Explaining the missing heritability of psychiatric disorders

Evidence from family, twin and adoption studies indicates that psychiatric disorders are substantially heritable. Heritability is usually expressed as the proportion of trait variance attributable to additive genetic factors (narrow sense heritability: h^2). The h^2 estimates for schizophrenia, attention-deficit/hyperactivity disorder, autism spectrum disorder and bipolar disorder are all >0.66, and are substantial for a range of other psychiatric conditions[1].

This evidence has motivated the application of increasingly sophisticated genomic approaches, including genome-wide association studies (GWAS) and next generation sequencing, that have identified a large number of genetic risk factors across a range of psychiatric conditions[2]. These studies revealed that psychiatric disorders are highly polygenic, with the major component of the heritability captured so far coming from common alleles (population frequency >0.01) detected in GWAS.

While this is extremely encouraging, and has set up an empirical platform upon which future progress towards precision psychiatry can be built[2], estimates of h^2 accounted for by the genetic variants identified in GWAS have always been substantially lower than the estimates of h^2 from family, twin and adoption studies. This shortfall is not a peculiarity of psychiatric disorders; it is also seen in many polygenic diseases and traits, and has been termed the "missing heritability".

Three main explanations for this missing heritability have been proposed[3,4]. First, it is possible that the estimates of h^2 from family, twin and adoption studies were inflated due to confounding factors such as shared environment. Second, estimates of h^2 from genomic studies may be deflated as they do not account for non-additive genetic effects such as dominance and gene-gene interactions. Finally, it may be the case that many risk alleles have simply not been identified by GWAS, either because their effects are too small or because they are too uncommon.

While all of these hypotheses remain plausible, the last one has received support from recent studies of polygenic traits and diseases, suggesting that many causal variants remain unidentified. In order to understand this, a brief explanation of GWAS is required. These studies involve genotyping single nucleotide polymorphisms (SNPs) that are common in the population (typically 500,000 - 1 million SNPs with a population frequency >5%). Because common SNPs tend to be correlated with their neighbours – a phenomenon known as linkage disequilibrium (LD) – the genotypes of additional SNPs can be inferred through a statistical process known as "imputation". This greatly increases the number of SNPs available to GWAS (typically >10 million SNPs with a population frequency >1%). When researchers seek associations in GWAS, they need to correct for the large number of statistical tests by taking a stringent threshold for statistical significance (known as genome-wide significance). This greatly reduces the occurrence of false positives, but at the expense of causing many real associations to be missed.

Early studies that revealed the missing heritability focused only on SNPs that met genome-wide significance. Subsequent studies have shown that more accurate and larger estimates of h^2 can be obtained by considering all available SNPs together, including imputed as well as directly genotyped SNPs, and by using data from reference samples that have undergone whole-genome sequencing (WGS) to allow better imputation of rare variants.

When these approaches are implemented, the proportion of h^2 that is captured increases to around one- to two-thirds of that expected in polygenic traits and diseases[4], with h^2 estimates for schizophrenia, bipolar disorder and autism being 0.23, 0.25 and 0.17, respectively[5]. This indicates that a proportion of the missing heritability was carried by SNPs that currently lie below the genome-wide significance threshold and also those that were insufficiently correlated with common SNPs to allow accurate imputation. It is, therefore, anticipated that the increased power of GWAS obtained from a substantial increase in both the number of common SNPs and the sample size will result in many more risk variants of small effect meeting genome-wide significance, as well as improving estimates of heritability[4].

However, the ability of common SNPs used in GWAS to capture the effects of variants with which they are in low LD is limited. The application of exome sequencing and WGS to complex disease cohorts has confirmed the presence in the human genome of a large number of rare genetic variants (defined as having a population frequency <1%). Importantly, these are not well correlated through LD with common SNPs and are therefore not accurately imputed in GWAS.

Recent work applying WGS to a large population cohort[6] has shown that estimates of heritability made using rare as well as common variants are much closer to those predicted from family studies for both height and body mass index, with much of the increase coming from SNPs that could not be accurately imputed from GWAS.

It is well recognized that, when compared to height and body mass index, many psychiatric disorders are under greater negative selection, and this is expected to result in a greater contribution from rare risk alleles. It is, therefore, plausible that rare genetic variants could be particularly relevant to psychiatric disorders, meaning that future WGS studies in large samples could prove to be particularly fruitful.

The prospect of large scale WGS studies in psychiatry is certainly exciting and will likely reveal much about genetic architecture and biology, as well as delivering better predictive tools. Short-read sequencing (SRS), based on compiling reads from <150bp segments, is currently the most widely used approach to WGS, because of its low cost and high throughput. It is particularly powerful in identifying rare single nucleotide variants and small insertion/deletions[7]. Robust approaches have been recently introduced to detect structural variants such as duplications, deletions, inversions, and other changes involving larger DNA segments (generally greater than 50-100 bases long) that are likely to be relevant to psychiatric disorders[8].

While SRS will undoubtedly be increasingly and fruitfully ap-

plied in psychiatric genomics in the coming years, it has limitations imposed by the fact that it works by stitching together short reads *in silico*. This means that there are regions of the genome which are difficult or impossible to read, such as those containing large structural variants, repetitive sequences, extreme guanine-cytosine content, or sequences with multiple homologous elements within the genome. This is sometimes known as the "dark genome".

There are now a number of long-read sequencing (LRS) platforms that allow the analysis of segments of the human genome up to 200kb, and these are capable of shining a light into the dark genome. Emerging studies using LRS are identifying larger, more harmful structural variants and long repetitive elements[7,9], both of which are candidates for involvement in psychiatric disorders.

Psychiatric genomics is a work in progress. GWAS have been hugely successful in identifying the role of multiple common variants, but recent work on missing heritability suggests a need to focus now on rare variants, and in the next few years we can expect studies based upon both SRS and LRS technologies to do this.

Fully characterizing the genetic architecture of psychiatric disorders is likely to improve polygenic risk prediction for both screening and stratification, allow a better understanding of the underlying biological mechanisms of disease, and broaden the landscape of pharmaceutical targets[2].

Michael J. Owen, Nigel M. Williams
MRC Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, UK

1. Sullivan PF, Daly MJ, O'Donovan M. Nat Rev Genet 2012;13:537-51.
2. Rees E, Owen MJ. Genome Med 2020;12:43.
3. Manolio TA, Collins FS, Cox NJ et al. Nature 2009;461:747-53.
4. Visscher PM, Wray NR, Zhang Q et al. Am J Hum Genet 2017;101:5-22.
5. Cross-Disorder Group of the Psychiatric Genomics Consortium. Nat Genet 2013;45:984-94.
6. Wainschtein P, Jain D, Yengo L et al. bioRxiv 2019;588020.
7. Chaisson MJP, Sanders AD, Zhao X et al. Nat Commun 2019;10:1784.
8. Collins R, Brand H, Karczewski K et al. bioRxiv 2019;578674.
9. Beyter D, Ingimundardottir H, Eggertsson H et al. Nature 2020;581:444-51.

DOI:10.1002/wps.20870

Toward a systems-based approach to understanding the role of the sympathetic nervous system in depression

The sympathetic nervous system (SNS) has an essential role in the prototypical stress response. Stress, stressors, and stress responses are central themes in most prominent theories of depression etiology and maintenance. Yet, the SNS is not a commonly targeted mechanism in depression research. Here we propose a dynamic, systems-level approach that contextualizes SNS-mediated stress responsivity within a regulatory framework. We believe that this conceptualization hews closer to the role of the SNS as a time-varying, context-driven regulatory system, and provides clinicians and researchers with a model for understanding its relevance to depression.

Interest in the SNS in depression is not new. A host of methods and markers have been used to try to delineate the role of the SNS in depression, including cardiac measures such as heart rate and pre-ejection period, skin conductance, salivary alpha-amylase, and urinary and serum measures of catecholamines. However, evidence for tonically-elevated SNS arousal in depression has been inconsistent and equivocal[1].

We propose three reasons for this equivocality. First, because the SNS is embedded in a larger set of regulatory systems, analysis of absolute levels should be augmented with – if not eschewed altogether for – a systems perspective that incorporates dynamic interrelations between system components. Second, the temporal dynamics of the stress response have been well documented[2], with SNS effects occurring relatively rapidly and ephemerally (compared to those of glucocorticoids), and attempts should be made to capture these time-dependent fluctuations. Third, there are likely to be individual differences in the dynamics and calibration of cognitive, affective and physiological regulatory systems. Thus, attempts should be made to identify subgroups.

Cognitive theories of depression have long posited the importance of depressogenic schemas – internal working models of the self, others, and the world – that magnify and distort the perception of ambiguous stimuli[3]. The presence of these schemas can increase the likelihood of threat appraisals (e.g., perceptions of external stressors) and the elicitation of negative emotional responses. The aversive arousal from negative emotions has been proposed to amplify memory for negative events[2] and provide experiential feedback that supports and reinforces the initial threat appraisal[3]. Thus, individuals with depression may be more likely to perceive environmental stressors, which elicit negative emotional reactions that reinforce the threatening nature of the stimulus and enhance memory encoding of the experience.

Inherent to this positive feedback loop between perceptions, appraisals and arousal is the physiological stress response to perceived stressors. This response serves an adaptive function to mobilize energy, stimulate immune activation, and increase cardiovascular tone through vasoconstriction and increases in heart rate and contractility. The stress response is composed of coordinated actions of the hypothalamic-pituitary-adrenal (HPA) axis, the SNS and the parasympathetic nervous system (PNS).

Compared to the SNS, there has been an abundant amount of research on the HPA axis and the PNS in depression, and studies have found evidence for HPA dysfunction[4] and reduced heart rate variability[1,5] in depressed patients. However, contradictory and null findings have also been common. We raise the possibility that inconsistent findings may stem from the isolation of system components in lieu of the whole. For instance, the HPA axis